Lecture30

Saturday, January 1, 2022   4:36 PM

Lecture30

Linear Algebra

Samira Hossein Ghorban
s.hosseinghorban@ipm.ir

Fall, 2021

Review: A useful lemma for self-adjoint matrices

$A A v = \lambda v \quad v \neq 0 \quad \langle Av, Av \rangle$

$v^* A^* A v = \lambda^2 v^* v$

Lemma

$A^* = A$

Let $A \in M_n(\mathbb{F})$ be a self-adjoint matrix. Then there is an invertible matrix $P \in M_n(\mathbb{F})$ such that $P^*P = I$ and

$$A = P \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} P^*,$$

where $\lambda_1, \dots, \lambda_n$ are eigenvalues of A.

- Singular value decomposition (SVD) is a generalization of this where $A \in M_{mn}(\mathbb{F})$ does not have to be self-adjoint or even square. No restriction at all!

$A = \begin{bmatrix} 1 & -1 & 3 & 5 & 0 \\ 3 & 0 & 4 & 6 & 7 \end{bmatrix}_{2 \times 5}$

$A \in M_n(\mathbb{F})$

$A = P D P^*$

جوردن ش

$\mathbb{F} = \mathbb{C}$

$A = LDU$

$A = S \begin{bmatrix} D \end{bmatrix} S^{-1}$

$A = S \begin{bmatrix} \; \end{bmatrix} S^{-1}$

AA^* and A^*A

- Let $A \in M_{mn}(\mathbb{F})$.
- The matrices $\underline{AA^*} \in M_m(\mathbb{F})$ and $\underline{A^*A} \in M_n(\mathbb{F})$ are self-adjoint and their eigenvalues are non-negative real numbers.
- There are unitary matrices $\underline{U} \in M_m(\mathbb{F})$ and $\underline{V} \in M_n(\mathbb{F})$ and diagonal matrices $D_1 \in M_m(\mathbb{R})$ and $D_2 \in M_n(\mathbb{R})$ such that

AA* مقادیر ویژه λ_1, λ_r

$$AA^* = UD_1U^* \qquad A^*A = V\underline{D_2}V^*.$$

$m\times m$ $\qquad$ $n\times n$

- We show that

$$D_1 = \begin{bmatrix} \lambda_1 & & & & & \\ & \ddots & & & & \\ & & \lambda_r & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{bmatrix} \quad D_2 = \begin{bmatrix} \lambda_1 & & & & & \\ & \ddots & & & & \\ & & \lambda_r & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{bmatrix}$$

$\underbrace{\qquad}_{m-r}$ $m\times m$ $\qquad$ $\underbrace{\qquad}_{n-r}$ $n\times n$

$A \in M_{mn}(\mathbb{F})$

$AA^* \in M_m(\mathbb{F})$

$mn \; nm$

$AA^* \in M_n(\mathbb{F})$

$(AA^*)^* = AA^*$

مقادیر ویژه A^*A

$\langle Ax, y\rangle = \langle x, By\rangle$

$\langle Ax, y\rangle = \langle x, A^*y\rangle$

$\langle Ax, x\rangle = \langle x, Ax\rangle$

$x^*\lambda x = x^*A^*x$, $\quad x^*Ax$

$\lambda = \bar{\lambda}$

Non-zero eigenvalues for AA^* and A^*A

$A \in M_{mn}(\mathbb{F})$

Lemma

Matrices AA^* and A^*A share the same non-zero eigenvalues with the same algebraic multiplicities.

$\lambda \neq 0$

$v \neq 0$ $\qquad$ AA^* مقدار ویژه λ

$$AA^*v = \lambda v \quad \Rightarrow \quad A^*A(A^*v) = \lambda A^*v$$

$\boxed{A^*v \neq 0}$

$A^*v = 0 \Rightarrow AA^*v = 0$

فرض کردیم مقدار ویژه غیر صفر $\lambda = 0$

$$\boxed{x^n f_{AB}(x) = x^m f_{BA}(x)}$$

در حالت کلی اگر $A, B \in M_{mn}(\mathbb{F})$ آنگاه

A, A^*

$$C = \begin{bmatrix} xI & A_{m\times n} \\ B_{n\times m} & I \end{bmatrix} \qquad D = \begin{bmatrix} I & 0 \\ -B & xI \end{bmatrix}$$

$\det CD = \det DC$

$$CD = \begin{bmatrix} xI & A \\ B & I \end{bmatrix}\begin{bmatrix} I & 0 \\ -B & xI \end{bmatrix} = \begin{bmatrix} xI - AB & xA \\ 0 & xI \end{bmatrix}$$

Sigular Values

- Suppose that $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_r > 0$ where $r = \operatorname{rank} A = \operatorname{rank} A^*A = \operatorname{rank} AA^*$.

$\operatorname{rank}(A) = \operatorname{rank}(A^*A)$

$\dim \underline{N(A^*A)} + \operatorname{rank}(A^*A) = n$

$N(A^*A) = N(A)$

- Let $\sigma_i = \sqrt{\lambda_i}$ for each $1 \leqslant i \leqslant r$, and consider m by n matrix

$$\Sigma = \begin{bmatrix} \sigma_1 & & & & \\ & \sigma_2 & & & \\ & & \ddots & & \\ & & & \sigma_r & \\ & & & & \\ & & & & \end{bmatrix}.$$

$m\times n$

$AA^* = U\underline{D_1}U^*$

$A^*A = VD_2V^*$

- We have

- We have
$$D_1 = \Sigma\Sigma^*, \qquad D_2 = \Sigma^*\Sigma.$$

$m\times n$

A candidate for a decomposition of A

- $AA^* = UD_1U^*$ $\qquad$ $A^*A = VD_2V^*.$ $\qquad$ $VV^* = I$
- $D_1 = \Sigma\Sigma^*$ $\qquad$ $D_2 = \Sigma^*\Sigma.$
- So, we obtain
$$AA^* = UD_1U^* = U\Sigma\Sigma^*U^* = U\Sigma \underbrace{V^*V}_{I}\Sigma^*U^* = \underbrace{(U\Sigma V^*)}_{A}\underbrace{(U\Sigma V^*)^*}_{A^*},$$
$$A^*A = VD_2V^* = V\Sigma^*\Sigma V^* = V\Sigma^*U^*U\Sigma V^* = \underbrace{(U\Sigma V^*)^*}_{A^*}\underbrace{(U\Sigma V^*)}_{A}.$$

λ_j $\quad Av_j$ $\quad v_j$

- It suggests that
$$A = U\Sigma V^*.$$

SVD

$\lambda_j = \sigma_j^2$

- Write V as $\begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix}$. For each $1 \leqslant j \leqslant r$,
$$A^*Av_j = \sigma_j^2 v_j.$$
- So,
$$AA^*(\underbrace{Av_j}_{\neq 0}) = \sigma_j^2 \underbrace{Av_j}_{\neq 0}.$$
- $\sigma_j^2 \neq 0$ implies that $Av_j \neq 0$, and consequently the unit eigenvector for eigenvalue σ_j^2 is $\frac{1}{\|Av_j\|}Av_j$.
- Write U as $\begin{vmatrix} u_1 & \cdots & u_m \end{vmatrix}$. So, $u_j = \frac{1}{\|Av_j\|}Av_j.$ $\Longrightarrow$ $u_j = \frac{1}{\sigma_j} Av_j$
- Also, $\|Av_j\|^2 = v_j^* \underbrace{A^*Av_j}_{\sigma_j v_j} = \sigma_j^2\|v_j\|^2 = \sigma_j^2$, so $\|Av_j\| = \sigma_j.$

A $\quad \begin{bmatrix} \lambda & & \\ & \ddots & \\ & & \lambda_r \end{bmatrix}_{m\times m}$

$m<n$ $\quad \Sigma = \begin{bmatrix} \sigma_1 & & & \\ & \ddots & & \\ & & \sigma_r & \end{bmatrix}$

$m>n$ $\quad \Sigma = \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_r \\ & & \end{bmatrix}$

$r \neq 0$ [illegible] λ

$A^*Av = \lambda v$ $\quad v = \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix}$

$v^*A^*Av = \lambda v^*v$

$\|Av\|^2 = \lambda \sum a_i^2$

$\lambda \geq 0$

$A = U\Sigma V^*$

$VV^* = I$

$\langle v_j, v_i \rangle = 0$

$A^*Av_j = \sigma_j^2 v_j$

A^*A [illegible] v_j

AA^* [illegible] Av_j

SVD

- Hence $Av_j = \sigma_j u_j$ for each $1 \leqslant j \leqslant r$.
- Consequently,

$A v_i = 0 \quad r+1 \le i \le n \quad V = [v_1 \cdots v_n]$

$v_1, \ldots, v_r, v_{r+1}, \ldots v_n$

$$
\begin{aligned}
AV &= A\begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix} \\
&= \begin{bmatrix} Av_1 & \cdots & Av_r & 0 & \cdots & 0 \end{bmatrix} \\
&= \begin{bmatrix} \sigma_1 u_1 & \cdots & \sigma_r u_r & 0 & \cdots & 0 \end{bmatrix} \\
&\quad \underbrace{\begin{bmatrix} u_1 & \cdots & u_r & 0 & \cdots & 0 \end{bmatrix}}_{U} \begin{bmatrix} \sigma_1 & & & & & \\ & \sigma_2 & & & & \\ & & \ddots & & & \\ & & & \sigma_r & & \\ & & & & 0 & \\ & & & & & \ddots \; 0 \end{bmatrix} \\
&= U\Sigma.
\end{aligned}
$$

$u_{r+1} \cdots u_m$

$AV = U\Sigma$

$\{u_1, \ldots u_r\} \subset \mathbb{R}^m$

$AV = U\Sigma$

$A = U\Sigma V^*$

SVD

$V = [v_1 \cdots v_n]$

This is "the " SVD

$$
\begin{aligned}
A &= U\Sigma V^* \\
&= \sigma_1 u_1 v_1^* + \cdots + \sigma_r u_r v_r^*
\end{aligned}
$$

$[u_1 \cdots u_r \; u_{r+1} \cdots u_m] \begin{bmatrix} \sigma_1 \ddots \sigma_r \ddots \end{bmatrix} \begin{bmatrix} v_1^* \\ \vdots \\ v_r^* \\ v_{r+1}^* \\ \vdots \\ v_n^* \end{bmatrix}$

Pseudo-inverse of A

$A \in M_{mn}(\mathbb{R}) \quad A = U\Sigma V^* \quad V\Sigma^\dagger U^*$

$N(A) \neq \{0\}$

$A: \mathbb{R}^n \xrightarrow{A} \mathbb{R}^m$

Assume

$$
\Sigma = \begin{bmatrix} \sigma_1 & & & \\ & \sigma_2 & & \\ & & \ddots & \\ & & & \sigma_r \end{bmatrix} \qquad \Sigma^\dagger = \begin{bmatrix} \frac{1}{\sigma_1} & & & & \\ & \frac{1}{\sigma_2} & & & \\ & & \ddots & & \\ & & & \frac{1}{\sigma_r} & \\ & & & & 0 \ddots 0 \end{bmatrix}
$$

The **pseudo-inverse** of A is defined as

$$A^\dagger = V\Sigma^\dagger U^*.$$

$A \cdot A^{-1} = I$

$A = U\Sigma V^*$

$AA^{-1} = I_n$

$A^{-1} = V\Sigma^{-1}U^*$

$m = n \neq r$

$m = n = r$

$\Sigma = \begin{bmatrix} \sigma_1 \ddots \sigma_r \ddots \end{bmatrix} \quad \Sigma = \begin{bmatrix} \sigma_1 \ddots \sigma_m \end{bmatrix}$

$\Sigma^\dagger = \begin{bmatrix} 1/\sigma_1 \; 1/\sigma_r \; 0 \end{bmatrix} \quad \Sigma^{-1} = \begin{bmatrix} 1/\sigma_1 \ddots 1/\sigma_n \end{bmatrix}$

The **pseudo-inverse** of A is defined as

$$A^\dagger = V\Sigma^\dagger U^*.$$

$A^\dagger$ fulfills the role of A^{-1}, "as far as possible."

$\Sigma^+ = \begin{bmatrix} 1/\sigma & & \\ & 1/\sigma & \\ & & \ddots \end{bmatrix}$ $\Sigma^{-1} = \begin{bmatrix} 1/\sigma_1 & & \\ & \ddots & \\ & & 1/\sigma_m \end{bmatrix}$

$A^+ = V\Sigma^+U^*$

For singular matrix A^*A

$A^*Ax = A^*b$

$\hat{x}$

- Let $x^\dagger := V^*\Sigma^+ Ub = A^\dagger b$;
- This $x^\dagger$ is a solution for $A^*Ax = A^*b$

$$\begin{aligned} A^*Ax^\dagger &= A^*AA^\dagger b \\ &= V\Sigma^*U^*U\Sigma V^*V\Sigma^\dagger U^*b \\ &= V\Sigma^*\Sigma\Sigma^\dagger U^*b \\ &= V\Sigma^*U^*b \\ &= A^*b. \end{aligned}$$

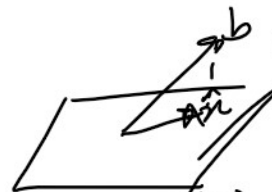
- Also, we show that $x^\dagger = V^*\Sigma^+ Ub = A^\dagger b \in \arg\min_x \|b - Ax\|$.

$Ax = b$

$x = A^{-1}b$

A^+

$Ax = b$ $\quad A \in M_{mn}(\mathbb{R})$



$E^2 = \|b - Ax\|^2$

$\hat{x} \in \arg\min E^2$

$C(A)$

$C(A)^\perp = N(A^*)$

$A^\dagger b$ is the optimal solution of $Ax = b$.

- The column space of $A^\dagger = V\Sigma^\dagger U^*$ is a sub-space of a spanning space generated by the first r columns of V.
- The null space of $A^*A = V\Sigma^*\Sigma V^*$ is equal to sub-space generating by the last $n - r$ columns of V.
- Since V is unitary,

$$C(A^\dagger) \perp N(A^*A).$$

- Assume that z is a solution of the equation system $A^*Az = A^*b$.
- We know that $A^*AA^\dagger b = A^*b$. So

$$A^*A(z - A^\dagger b) = 0 \Rightarrow z - A^\dagger b \in N(A^*A).$$

- Let $v = z - A^\dagger b$. Thus $z = v + A^\dagger b$ where $v \in N(A^*A)$ and $A^\dagger b \in C(A^\dagger)$.
- So, $\|z\|^2 = \|v\|^2 + \|A^\dagger b\|^2 \geqslant \|A^\dagger b\|^2 = \|x^+\|^2$

$A^+ = V\Sigma^{-1}U^*$

$C(A^T) \subseteq \mathrm{span}(\{v_1, \dots, v_r\})$

$A^*A = V\Sigma^*\Sigma V^*$ $\Rightarrow$ $N(A^*A) = \mathrm{Span}\{\{v_{r+1}, \dots, v_n\}\}$

$x^+ = A^+b$

$\|z\|^2 \geq \|x^+\|^2$

$z = v + A^+b$

$N(A^*A)$ $\quad C(A^+)$

Thank You!